

WHAT IS CLAIMED IS:

1. A method for classifying email messages, the method comprising
using a plurality of modules to determine a level of sameness of a particular email message with one or more prior email messages, wherein the level of sameness is derived for the particular email message from a weighting of the outputs of the modules;
determining a performance level for each of the modules;
comparing performance levels;
adjusting a weighting of at least one module in response to comparing performance levels; and
using the level of sameness for the particular email message to classify the particular email message into a category.
2. The method of claim 1, further comprising
comparing the number of email messages classified in the category with a predetermined number; and
if the number of email messages is greater than the predetermined number then
classifying the category as a first category type; else
classifying the category as a second category type.
3. The method of claim 2, wherein the first category type is bulk email.
4. The method of claim 2, further comprising
accepting a signal from a user input device to indicate processing of email messages in a category.
5. The method of claim 4, wherein the processing includes preventing the email messages in a category from being delivered to a user.
7. The method of claim 1, wherein a category is commercial email.
8. The method of claim 1, wherein Bayesian analysis is used.
9. The method of claim 1, further comprising

accepting a signal from a user input device to set a parameter; and
using the parameter to adjust a weighting.

10. The method of claim 1, wherein a module analyzes word count in an email message.

11. The method of claim 1, wherein a module analyzes similarity of text in an email message.

12. The method of claim 1, wherein a module analyzes a similarity of sender addresses.

13. The method of claim 1, wherein a module analyzes a similarity of network routing.

14. The method of claim 1, wherein a module uses a hash of information in an email message.

15. The method of claim 1, wherein a message classification in a bulk category includes a determination of whether the number of email messages in a category exceed a predetermined number, the method further comprising
submitting email messages in the bulk category to analysis to determine a level of commercial text.

16. The method of claim 15, further comprising
preventing messages with a predetermined level of commercial text from being sent to an intended recipient.

17. The method of claim 15, further comprising
intercepting email messages from being sent to an intended recipient;
collecting the intercepted messages for a period of time;
determining whether the collected messages are bulk messages, and if so, submitting the email messages in the bulk category to analysis to determine a level of commercial text.

18. The method of claim 17, further comprising preventing messages with a predetermined level of commercial text from being sent to an intended recipient.

19. The method of claim 1, further comprising assigning a lower rating to a module with a low performance level.

20. The method of claim 1, further comprising assigning a higher rating to a module with a high performance level.

21. The method of claim 1, further comprising preventing a module with a low performance level from being used in a subsequent determination of a level of sameness.

22. An apparatus for classifying email messages, the apparatus comprising a processor for executing instructions included in a machine-readable medium, the machine-readable medium including one or more instructions for using a plurality of modules to determine a level of sameness of a particular email message with one or more prior email messages, wherein the level of sameness is derived for the particular email message from a weighting of the outputs of the modules;

one or more instructions for determining a performance level for each of the modules;

one or more instructions for comparing performance levels;

one or more instructions for adjusting a weighting of at least one module in response to comparing performance levels; and

one or more instructions for using the level of sameness for the particular email message to classify the particular email message into a category.

23. A machine-readable medium including instructions executable by a processor for classifying email messages, the machine-readable medium including

one or more instructions for using a plurality of modules to determine a level of sameness of a particular email message with one or more prior email messages, wherein the level of sameness is derived for the particular email message from a weighting of the outputs of the modules;

one or more instructions for determining a performance level for each of the modules;
one or more instructions for comparing performance levels;
one or more instructions for adjusting a weighting of at least one module in response to comparing performance levels; and

one or more instructions for using the level of sameness for the particular email message to classify the particular email message into a category.

24. An apparatus for classifying email messages, the apparatus comprising means for using a plurality of modules to determine a level of sameness of a particular email message with one or more prior email messages, wherein the level of sameness is derived for the particular email message from a weighting of the outputs of the modules;

means for determining a performance level for each of the modules;
means for comparing performance levels;
means for adjusting a weighting of at least one module in response to comparing performance levels; and

means for using the level of sameness for the particular email message to classify the particular email message into a category.